IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

GOOGLE PLAY STORE APP RATING PREDICTION

Guide: Dr B Srinivasa Rao

Professor

Gokaraju Rangaraju Institute of Engineering and Technology

Authors:

M Sai Kiran, P V S Karthikeya, M S Kushal Raj, Manuka Varun Teja

Student, CSE,	Student, CSE,	Student, CSE,	Student CSE

Gokaraju Rangaraju Institute of Engineering and Technology

***Abstract**:* In order to create new software, it is necessary to adhere to established implementation standards. The development of applications for marketplaces has been difficult when it comes to the sale and acceptance of software by clients. The number of downloads, the number of comments, and the number of ratings on apps can all be found in app stores. A pattern of app success may be deduced from the obstacles encountered in these sectors and their methods of problem analysis. It was intended to construct two inference engines utilizing SVM and Random Forest algorithms based on this scenario, as well as to research the attributes that best correlate with app ratings and calculate and assess regression metrics using a Google Play Store database.

Google Play Store Apps, Predicting Ratings, and Machine Learning are some of the keywords.

1. INTRODUCTION

Many of the problems we face can only be solved via machine learning techniques. Machine learning models and architectures are described in great detail in this work. There is a plethora of uses for machine learning, as well as several potential directions in which it can go in the future. Speculations based on machine learning are expected to become more commonplace in the future. In the meantime, its unsupervised learning capabilities will be increased because there is a lot of information out there, but it isn't necessary to name every one of them. It is also expected that brain system topologies will become more unpredictable in order to separate the most semantically significant highlights from the less important ones. In addition, we'll be able to use the advantages of deep learning and improved adaptability to complete additional assignments.

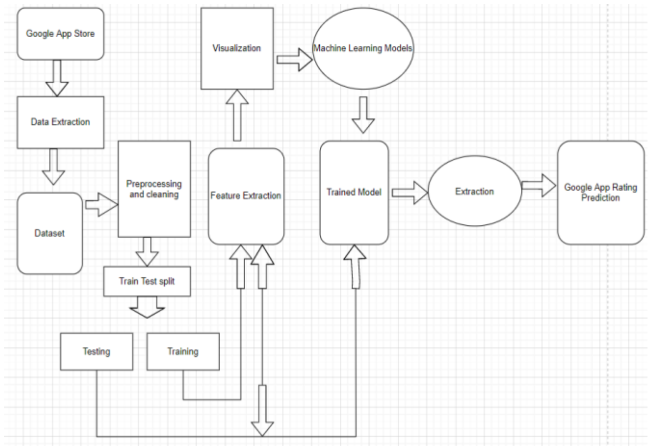

Fig.1: Architecture

Subjective data can be found and recognized using methods such as opinion mining that combine natural language processing with text analysis and computational linguistics. In this subfield of text mining, algorithms are being developed to locate and extract textual representations of people's ideas and thoughts. People's own beliefs and values play a significant role in the decisions they make on a daily basis. Smartphones are having an ever-increasing impact on our personal and professional life. Smartphones and other mobile devices now have applications (apps) that provide services including medical care, fitness, beauty, monitoring, and sports. [1] As shown in Figure 1A, there were around 2.6 million apps accessible on the Google Play store as of March 2019. Numerous users will be able to submit feedback in the form of user reviews and numerical ratings for these apps. Figure 1B shows that there are currently 205.4 billion app downloads per year, with a projected increase to 352.9 billion by 2021. The most significant elements for other users are user reviews and numerical ratings [2]. User reviews and numerical ratings have a considerable impact on the general adoption of mobile applications, according to research [3]. Buying a 5-star rated goods will cost you 20–99 percent more than a 4-star rated product, according to study. Shared ratings, problem reports, and reviews benefit both app users and app developers [4].

These problems can be addressed in several ways [6, 7], but all of them rely on just seeing the extremes in user reviews' polarity. Polarity and subjectivity can be described in three ways: positive, negative, or neutral. There is a problem with these solutions because they only take into account user reviews and not real app ratings, neglecting the issue of numerically out of sync ratings. As a result of our studies, we've discovered that the numerical ratings and user reviews that appear on the product's page are frequently inconsistent.

2. LITERATURE REVIEW

2.1 User feedback in the Appstore: an empirical study [4]:

Customers can rate and review applications they've downloaded from app stores like Google Play and the Apple App Store. Over the past few years, these platforms have become increasingly popular, attracting both app developers and end users. To what extent these changes will alter the way requirements engineering activities are carried out remains an open question. Apple AppStore review data was analyzed in this study, which included over one million reviews. It was important to see how and when users offered feedback, and how the content of the feedback influenced the community as a whole. Most of the feedback occurs in the first few days after a new release, but then levels off rapidly. Users' experiences, problems they've encountered, and feature requests are some of the most common topics reviewed by reviewers. In terms of utility and creativity, there is certainly a wide range of offensiveness. Higher ratings are a direct outcome of positive feedback, and vice versa is also true when it comes to the number of downloads. Many people have a tendency to focus on the bad parts of a situation, such as criticism, rather than the context and user experience. Our findings have a substantial impact on software development and requirements engineering.

2.2 Analysis of social media comments from Thai customers [5] using sentiment analysis:

Recent years have seen an increase in the importance of sentiment analysis technology (also known as opinion mining) with regard to the growing volume of Thai online consumer reviews available on social media and websites. This method analyses people's feelings, views, attitudes, and sentiments. Using a "bag of words" technique for opinion mining, a variety of keywords can be utilized to distinguish between positive and negative input.

Because of their dispersed nature, European language evaluations are well-suited to these approaches.

These strategies relying on a jumble of words have a problem when it comes to customer reviews from Thai customers. These approaches cannot be employed because Thai texts are composed of a long succession of characters without word boundaries. There hasn't been much research done on Thai customer sentiment analysis up to this point.

2.3 An algorithm based on user evaluations is used to identify features in mobile apps:

Google Play's store rating and user input have a significant impact on App Store Optimization. An app's overall quality can be judged based on user feedback and the average number of stars awarded to it. We don't trust the star rating as a measure of consumer satisfaction because it's just a phrase in mathematics. A single number on a scale from one to five will not give us a complete picture of how people feel about the app. As a result, we advocate a different methodology to evaluating app features based on user feedback, which we believe delivers a more accurate app rating. Reviewers' comments were mined for popular aspects as well as for evaluation. To begin, we scoured the Google Play Store for reviews from various product categories. We were able to establish the degree of agreement and disagreement among the assessments using sentiment analysis. Each evaluation was given a polarity strength rating.

We averaged the ratings of each reviewer to get the final score. We then compared the user-generated star ratings to the established star ratings. Existing ratings were found to be substantially higher than the aggregate review scores. For the most popular features, machine learning was utilized to determine their popularity, which was based on the app's category. We came up with app features and their popularity based on user feedback. Use these rankings to understand about the most popular features and why they are so popular with potential app users. Until then, app developers can assess which functionalities are being underutilized and seek to improve those features.

2.4 To determine whether or not popular free hybrid Android/iOS apps maintain their user ratings and reviews over time:

The major purpose of these hybrid development tools is to provide an app that is regarded the same by consumers regardless of platform. When deciding whether or not to download an app, customers have previously relied on star ratings and user reviews. Given the importance of star ratings and user reviews, we study whether the adoption of a hybrid app development tool helps developers achieve consistency in the star ratings and user reviews across several platforms. We examined 68 "hybrid apps," or programs available on both Google Play and the Apple App Store, for this investigation. Inconsistent star ratings were detected on 33 of the 68 hybrid apps we tested.

2.5 Sentiment analysis and opinion mining in an e-commerce application [8]:

Many people use social media to communicate their personal perspectives on current events and other issues of common interest. This information can be helpful when making decisions. Before making a decision, people and businesses can learn about other people's perspectives on a topic by using social media. [9] Study finds that businesses all over the world are discovering that e-commerce isn't only for buying and selling products online, but also for enhancing operational efficiency to compete with the market's other big players. User choice in topics, peer influence, and user profile information all have an effect on what they think about a certain subject.

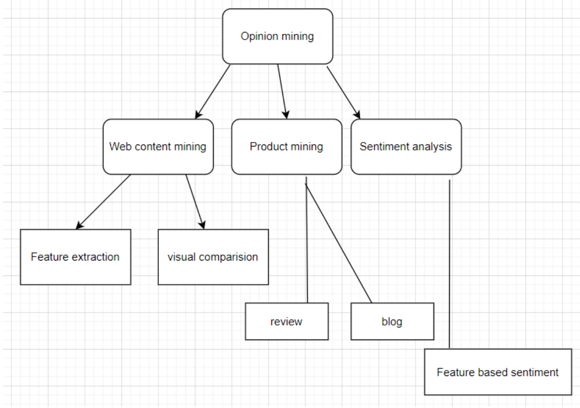

Fig.2: Analysis of Opinion Mining

2.6 Aspect analysis for opinion mining of Vietnamese text [10]:

Opinion mining's most difficult task is aspect extraction. Many studies have been conducted to find a solution to this issue in English text. Languages like Vietnamese, which aren't as widely spoken, have a more difficult time managing and analyzing their complicated structures. Semi-supervised GK-LDA outperforms classical topic modeling LDA in terms of performance. We employ a dictionary-based strategy to extract noun-phrases for better performance than only extracting word seeds or using a complete sentence to infer aspects in the aspect inference. The results of our tests suggest that the approach we've described works well for extracting and classifying aspects. Our method is based on Vietnamese literature, but we believe it may be applied to other languages as well.

3. IMPLEMENTATION

Random Forest and Support Vector Machine algorithm are used in this project for better accuracy and performance than current app rating criteria. Classification and Regression are two of the most commonly used machine learning algorithms, and Random Forest is the most commonly used algorithm for this purpose. It is possible to use the Support Vector Machine for both classification and regression challenges because it is a supported machine learning algorithm.

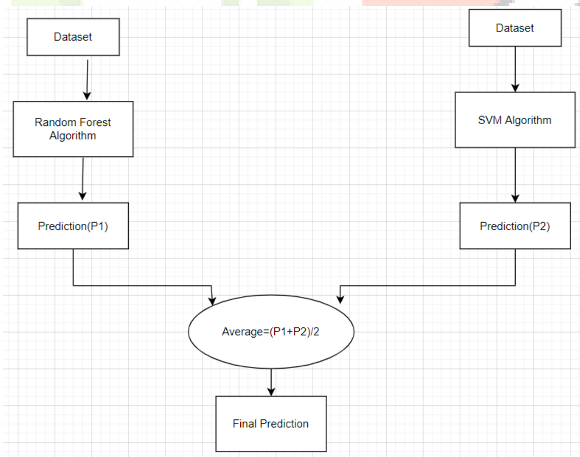

Fig.3: Flow Chart

Analysis and Prediction:

The importance of mobile apps in the lives of today's consumers cannot be overstated. The growth of mobile app advertising has been shown to have a significant impact on cutting-edge innovation. [11] Nevertheless, as the market for mobile apps grows, so does the number of mobile app designers, resulting in a rise in the global mobile app industry's earnings that can only be described as spectacular. In the face of great competition from all over the world, it is essential for a designer to recognize that he is moving in the right direction.

The application designers may have to find out how to maintain their current position in the market in order to retain this money and their position in the market. According to reports, the Google Play Store is the most popular software distribution channel. Despite the fact that it generates more than double the number of downloads as the Apple App Store, it only generates a small fraction of the revenue. This is how we scraped data from the Play Store in order to focus our investigation there.

Advanced cell technology has made portable applications (Mobile Apps) an essential part of our daily life. [12] In spite of the fact that new apps are being released every day, it is difficult for us to keep up and keep up with the facts and to comprehend everything about the apps. In September 2011, Android1market had over a million apps, which can be verified. As of right present, the Google Play App Store offers 0.675 million Android applications. Having such a large number of applications gives customers the opportunity to choose from a wide range of options.

It is our belief that paid applications are influenced by online application questionnaires. For a potential customer, perusing all of the literary comments and ratings can be a challenge. As a result, application engineers have difficulty figuring out how to improve the program's performance based solely on general evaluations, and would benefit from reading through the many printed comments.

Automated Intelligence: Automated systems can learn and develop without being explicitly programmed through the use of machine learning, a form of artificial intelligence (AI). Automated learning is concerned with the possibility of computer programs being able to access data and make their own decisions based on that data.

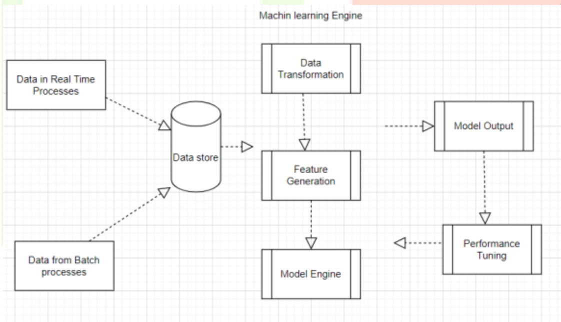

Fig.4: Machine Learning Architecture

When training a computer how to learn from a dataset or other information, the phrase "supervised learning" is used to describe the process. After analyzing the data (models) provided, sustained learning generates a right response from the input provided [13].

Unsupervised learning does not make use of any current data or input throughout the training process. The lack of a teacher means that there are no students to instruct. Instead of telling the algorithm what to do, we leave it to it. Without any prior training or direction, the machine's primary job is to find patterns and similarities in the incoming dataset.

As a result of being just partially monitored:

Supervised and unsupervised methods are used in semi-supervised machine learning. A "labeled" dataset is used to train a machine learning algorithm using more common methods of supervised machine learning.

4. PROPOSED MODEL

In this project, we use Random Forest and Support Vector Machine Algorithm for better accuracy and performance as compared to existing app rating predictions. Random forest is the most used supervised machine learning algorithm for Classification and Regression (it predicts categorical and continuous outcome). Support Vector Machine is a supervised machine learning algorithm that can be used for both classification or regression challenges.

In this project we can implement a model by following steps:

1. Collect data and import the suitable libraries to the platform called Jupyter Notebook(anaconda3) using Python language.

2. Analyzing data by graphical representation such as decision trees, Bar graphs, and Histograms.

3. Data wrangling is used to clean the data by removing unwanted columns from the dataset.

4. We build model on the train data and predict the output on the test data.

5. Accuracy check: we can find the actual and predicted values by using existing methods called root mean square error, r-square, non-linear equation, and mean absolute error.

6. Results of Random Forest and SVM algorithms, we collect the outputs of and find the averages for better accuracy and performance of google play store app rating prediction.

5. METHODOLOGY

SVM:

Support Vector Machine (SVM) is a supervised machine learning algorithm used for both classification and regression. Though we say regression problems as well its best suited for classification. The objective of SVM algorithm is to find a hyperplane in an N-dimensional space that distinctly classifies the data points. The dimension of the hyperplane depends upon the number of features. If the number of input features is two, then the hyperplane is just a line. If the number of input features is three, then the hyperplane becomes a 2-D plane. It becomes difficult to imagine when the number of features exceeds three. Let's consider two independent variables x1, x2 and one dependent variable which is either a blue circle or a red circle.

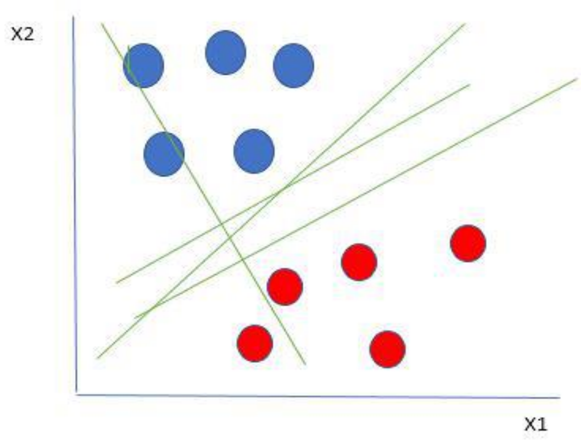

Fig.5: SVM model

From the figure above its very clear that there are multiple lines (our hyperplane here is a line because we are considering only two input features x1, x2) that segregates our data points or does a classification between red and blue circles. So how do we choose the best line or in general the most effective hyperplane that segregates our data points?

SVM Kernel:

The SVM kernel could be operated that takes low dimensional input space and transforms it into higher-dimensional space, it converts not divisible problem to separable problem. It's largely useful in non-linear separation problems. Simply put the kernel, it does some extremely complex data transformations and then finds out the method to separate the data based on the labels or outputs defined.

Advantages of SVM:

- Effective in high dimensional cases
- Its memory-efficient because it uses a subset of training points in the decision function called support vectors
- Different kernel functions can be specified for the decision functions and its doable to specify custom kernels

Types of SVM:

SVM is of two types:

Linear SVM: Linear SVM is employed for linearly separable data, which means if a dataset is classified into 2 categories by employing a single straight line, then such data is termed linearly separable data, and classifier is used called Linear SVM classifier.

Non-linear SVM: Non-Linear SVM is employed for non-linearly separated data, which suggests if a dataset can't be classified by employing a straight line, then such data is termed non-linear data, and the classifier used is named as Non-linear SVM classifier.

RANDOM FOREST:

Random Forest could be a machine learning algorithm that belongs to the supervised learning technique. It is used for each Classification and Regression problem in ML. It is based on the concept of ensemble learning, which could be a method of mixing multiple classifiers to solve a complex problem and boost the performance of the model. because the name suggests, "Random Forest could be a classifier that contains a variety of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset." Rather than hoping on one decision tree, the random forest takes the prediction from each tree, and based on the majority votes of predictions, it predicts the final output. The bigger number of trees in the forest leads to higher accuracy and prevents the matter of overfitting.

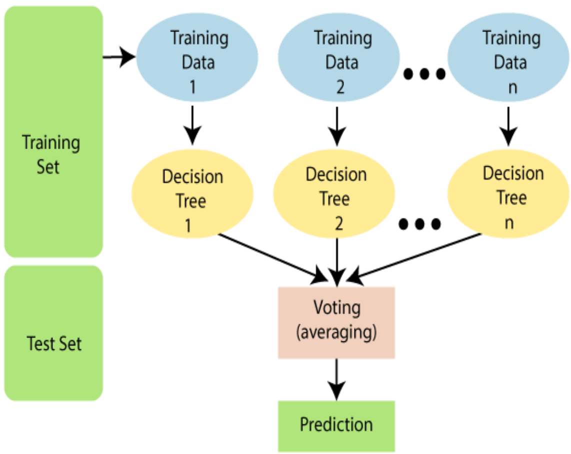

Fig.6: Random Forest model

Since the random forest combines multiple trees to predict the category of the dataset, it's doable that some decision trees could predict the proper output, while others may not. But together, all the trees predict the correct output. Therefore, below are 2 assumptions for a better Random Forest classifier:

There should be some actual values within the feature variable of the dataset in order that the classifier will predict correct results instead of a guessed result.

The predictions from each tree must have very low correlations.

Random Forest works in two-phase 1st is to form the random forest by combining the N decision tree, and 2nd is to form predictions for every tree created in the 1st phase.

The operating method can be explained within the below steps:

Step 1: Build the Support vector regressor model.

Step 2: Build the Random Forest Regressor.

Step 3: Predict the rating using both models.

Step 4: Calculate the average of both models.

Step 6: Our model is finished.

5. EXPERIMENTAL RESULTS

Preprocessing will take away null values and duplicates from a reliable dataset.

The dataset is separated into 2 sections: a training dataset and a testing dataset using the train test split function.

A training dataset is needed for the development of machine learning models.

This approach makes use of machine learning models that enable vector regression and random forest regression.

The test data set will be used to train models, which can then be used to create predictions.

Based on the results of the 2 machine learning approaches, the final projection is calculated.

MODULE 8: PREDICTION ON DATA FEATURES

In [35]:

```
apps_in_test_set = apps_in_test_set.reshape(-1, 1)
avg = avg.reshape(-1, 1)
y_test = y_test.reshape(-1, 1)
average = np.concatenate([apps_in_test_set, avg,y_test], axis = 1)
average = pd.DataFrame(average, columns = ['App Name', 'Predicted Rating','Actual Rating'])
average
```

Out[35]:

	App Name	Predicted Rating	Actual Rating
0	Tattoo Photo Editor- Photo Tattoos, Tattoo Maker	4.0	4.1
1	Handcent 6 ThanksGiving Skin	3.6	4.3
2	Monkey Sounds	4.4	3.8
3	HealthSpective	3.9	1.7
4	Craft King	3.6	4.2
...	...	...	...
119995	Massage Guide: Learn Relaxing Massages	3.8	2.8
119996	Kamus Obat Terbaru (Lengkap & Praktis)	3.8	4.4
119997	Interprétation des rêves - Signification	4.3	4.2
119998	Loco Craft: 3 Creative Maps	4.0	4.1
119999	Save Status to Gallery And Was tickers Text	4.0	4.1

120000 rows × 3 columns

Fig.7: Prediction screen

Random Forest Regression:

Mean absolute error: 0.540935

Mean squared error: 0.518304

Root mean squared error: 0.719936

Support Vector Regression:

Mean absolute error: 0.551403

Mean squared error: 0.597192

Root mean squared error: 0.772781

Combined Model:

Mean absolute error: 0.520076

Mean squared error: 0.488790

Root mean squared error: 0.699135

6. CONCLUSION

We came to the conclusion that our hypothesis is correct after running through all of these algorithms and processes. As a result, it is possible to predict app ratings, but a large amount of preprocessing is required before the classification and regression processes can be started. The data collected from Google Play Store apps has huge potential to help app development companies succeed. Developers can use the information to their advantage to work on and conquer the Android market! In order to accurately estimate whether an app will have more than 100,000 downloads and be a success on the Google Play Store, we need to know the app's Size, Type, Price, Content Rating, and Genre.

7. FUTURE SCOPE

Only polarity and subjectiveness may be obtained from user reviews. Predictions are also important because of the enormous expansion in review-based data. This is a challenging but rewarding process, as user reviews are qualitative and ratings are mainly quantitative. Additionally, Google's numerical rating system may be distorted and amplified by the fact that higher ratings supplied by consumers may bring in disproportionately more new users. So this study investigated if ensemble classifiers might be used to predict numerical ratings for Google Play store apps based on user reviews. The Google App store assessments were used to test many ensemble classifiers. To predict numerical ratings in the future, deep learning technology will be applied

REFERENCES

[1] Statista, Numerous accessible applications within the Google Play store from December 2009 to March 2019, https://www.statista.com/statistics/266210/numberof-available-applications-in-the-google-play-store/, Online: accessed twenty-two May 2019.

[2] Statista, various mobile app downloads worldwide in 2017, 2018, and 2020 (in billions), https://www.statista.com/statistics/ 271644/worldwide-free-and-paid-mobile-app-store downloads/, Online: accessed twenty-two May 2019.

[3] Online shopping, pew internet, and American life project, Washington, DC, 2018, http://www.pewinternet.org/Repor ts/2008/online shopping/01-Summary-of-Findings.aspx Online: accessed eight August. 2014.

[4] D. Pagano and W. Maalej, User feedback within the AppStore: an empirical study, in Proc. IEEE Int. Requirements Eng. Conf. (Rio de Janeiro, Brazil), July 2013, pp. 125–134.

[5] T. Chumwatana, Using sentiment analysis technique for analyzing Thai client satisfaction from social media, 2015.

[6] T. Thiviya et al., Mobile apps feature extraction supported user reviews using machine learning, 2019.

[7] H. Hanyang et al., learning the consistency of star ratings and reviews of standard free hybrid android and ios apps, Empirical Softw. Eng. 24 (2019), no. 7, 7–32.

[8] N. Kumari and S. Narayan Singh, Sentiment analysis on e-commerce application by using opinion mining, in Proc. Int. Conf.-Cloud Syst. Big Data Eng. (Noida, India), Jan. 2016, pp. 320–325.

[9] R. M. Duwairi and that I. Qarqaz, Arabic sentiment analysis using supervised classification, in Proc. Int. Conf. Future Internet Things Cloud (Barcelona, Spain), Aug. 2014, pp. 579–583.

[10] H. S. Le, T. V. Le, and T. V. Pham, Aspect analysis for opinion mining of vietnamese text, in Proc. Int. Conf. Adv. Comput. Applicant. (Ho Chi Minh, Vietnam), Nov. 2015, pp. 118–123.

[11] H. Wang, L. Yue, and C. Zhai, Latent aspect rating analysis on review text data: a rating regression approach, in Proc. ACM SIGKDD Int. Conf. Knowledge Discovery Data Mining (Washington, D.C., USA), July 2010, pp. 783–792.

[12] K. Dave, S. Lawrence, and D. M. Pennock Mining the peanut gallery: Opinion extraction and semantic classification of product reviews, in Proc. Int. Conf. World Wide Web (New York, USA), 2003, pp. 519–528.

[13] B. Pang, L. Lee, S. Vaithyanathan, Thumbs up?: sentiment classification using machine learning techniques, in Proc. ACL-02 Conf. Empirical MethodsNatural Language method. (Stroudsburg, PA, USA), 2002, pp. 79–86.